

of coverage, a heterogeneity that was further magnified by measuring the criterion on a cash or accrual basis, and in above- or below-the-line terms. This wide variety of concepts often carried over to the data reported in the *WEO*, thus putting paid to the notion that the numbers included in *WEO* are strictly comparable.<sup>29,30</sup>

**... highlighting the importance of having countries adopt standard data templates and—in the meantime—providing clear metadata.**

32. These findings underscore the importance of providing clear metadata for all IMF-disseminated data. However, notwithstanding the IMF's guidelines to staff to this effect, an examination of a large sample of Article IV consultation reports showed that none of them complied with this requirement (*Annex 5*). An inattentive or impatient economist could simply download the desired data, compare apples and oranges, and draw the wrong policy conclusions.<sup>31</sup> Nevertheless, even with excellent metadata, the diversity of definitions can greatly impair the ability to do cross-country work.

### (iii) Financial surveillance

**Data issues for financial surveillance are among the most challenging ...**

33. Financial sector surveillance, in the aftermath of the global financial crisis, has become even more central to the Fund's core operations. However, data issues are particularly challenging here, given the sensitive (and often confidential) nature of the data, the need for granularity and comprehensiveness (e.g., “off-balance-sheet” exposures), and lack of consistency. Data are often

<sup>29</sup>Nominal GDP provides another example of comparability issues in *WEO* data. While most countries still measure GDP using the 1993 System of National Accounts (SNA), some, including most of the advanced economies, have now moved to the 2008 SNA. Typically, GDP, as measured under the 2008 SNA, is larger than that under the older system (e.g., U.S. nominal GDP was almost 4 percent larger, while it is estimated that, were China to move to the newer system, its economic size could be as much as 16 percent larger).

<sup>30</sup>The *WEO* makes adjustments to some data to improve comparability. For example, the *WEO* has migrated balance of payments data to the methodology used in the sixth edition of the *Balance of Payments and International Investment Position Manual (BPM6)*, even though many countries still submit data under the previous *BPM5* methodology. The *WEO* also converts data on a fiscal-year basis to a calendar-year basis.

<sup>31</sup>In interviews with external data users, many admitted that they use multiple (noncomparable) IMF data sources (*IFS*, *WEO*, country reports, Working Papers) to fill in missing data for cross-country studies.

nonexistent or opaque in some critical areas, particularly on cross-border linkages and the shadow banking sector.

**... due, in large part, to the often market-sensitive nature and need for granularity of data.**

34. Financial surveillance is constantly struggling with the tension between granularity and aggregation. Aggregate data can mask critical vulnerabilities—that granular data might reveal—and may not be usable with some of the Fund's new analytical tools. For example, network analysis (used to examine issues of interconnectedness) needs quite granular data. But the Fund's hands are essentially tied by its Articles of Agreement, as it cannot require countries to provide institution-specific data.

35. The Financial Stability Assessment (FSA), a component of the FSAP and a key instrument of the IMF's surveillance, illustrates the data challenges facing the IMF. According to this evaluation's survey of staff, the data collected for FSAPs are perceived as the most problematic.<sup>32</sup> Just under a third of the survey respondents from the Monetary and Capital Markets Department (MCM) believed data were sufficient for conducting an FSAP exercise. Notably, almost 90 percent said data problems had hindered the conduct of stress tests, while about three-quarters said the analysis of potential cross-border spillovers was hampered by data problems. Poor quality data was cited by 40 percent of respondents.

36. A 2014 review of the FSAP (IMF, 2014d) made clear the role of data—in particular, availability and quality for stress testing—in underpinning (or undermining) the program's effectiveness. While many countries voluntarily provide these data to the FSAP team—subject to strict confidentiality protocols—the FSAP review noted that this practice is not universal (as confirmed by this evaluation's survey results), with advanced countries the least likely to share supervisory data. And even when the needed data are available, FSAP teams are typically not equipped to assess their accuracy or the quality of underlying assets.

**... underscoring the importance of building trust, yet being candid about data limitations.**

37. Considering the reasons why country authorities are loath to share data, about a third of MCM survey

<sup>32</sup>Also, compared with staff working on the *WEO*, staff involved with the *Global Financial Stability Report* were much more likely to note problems with lack of data, comparability, and uncertain quality.

respondents cited legal constraints,<sup>33</sup> but another third cited issues of trust. Data providers who were interviewed regarding access to market-sensitive financial data noted that banking supervisors tend to trust, in order, other supervisors, central banks, the Bank for International Settlements (BIS), and only then the IMF.<sup>34</sup> Further, as noted by the 2014 Triennial Surveillance Review, “some of the Fund’s counterparts have become less willing to share data as the crisis has subsided.” This is clearly problematic for the Fund’s FSAP, which in such cases must rely on publicly available data and/or on stress tests conducted by supervisors and the banks themselves. According to some interviewees, the results from the FSAP team’s stress tests differed at times from those of the stress tests conducted by the authorities or the banks, largely because of differing access to data.

38. These findings highlight two issues: (i) there is a tension between the mandatory character of FSAs and the voluntary provision of the data they require,<sup>35</sup> and (ii) the limitations of the associated risk assessment need to be clearly communicated by the FSAP team. The evaluation survey of MCM staff is revealing in this regard, with 40 percent of respondents advocating mandatory data provision to help address data deficiencies, and only half agreeing that the Financial System Stability Assessment (FSSA) report had clearly noted the problems with data quality or access. To help address the first issue, the Fund could clarify its confidentiality protocols to the membership to encourage the voluntary provision of the needed information.<sup>36</sup> On the second issue, the 2014 FSAP review noted that the standard disclaimer on all FSSA reports should be expanded to highlight any data limitations. But this evaluation found no change in the standard disclaimer in the most recent FSSA reports, including some with serious data access and quality issues.

39. The Fund has made notable strides, nonetheless, with respect to data needed for financial surveillance.

<sup>33</sup>Some countries with legal constraints find ways to allow the FSAP team to “access” the data without actually violating the law (e.g., letting the FSAP team into the room to watch the conduct of supervisory stress tests).

<sup>34</sup>The 2013 IEO evaluation, *The Role of the IMF as Trusted Advisor*, also found that country authorities placed more trust in the BIS than the Fund in the handling of confidential data (IEO, 2013).

<sup>35</sup>In September 2010, the Executive Board decided to make the Financial Stability Assessment (FSA) mandatory for systemically important financial sectors in response to the shortcomings revealed by the financial crisis. Previously, all FSAs, as part of an FSAP exercise, were conducted on a strictly voluntary basis.

<sup>36</sup>See “Confidentiality Protocol—Protection of Sensitive Information in the Financial Sector Assessment Program,” IMF, *Selected Decisions*, Thirty-Second Issue, p. 108.

Substantive progress has been made since the global crisis on collecting data on Globally Systemically Important Financial Institutions (G-SIFIs), the nonbank financial institutions (NBFIs), shadow banking, and Financial Soundness Indicators (FSIs).<sup>37, 38</sup> The Fund’s efforts to collect data on NBFIs and shadow banking is particularly important to allow the Fund to expand its coverage of stress tests to the nonbank sector (an increasingly important player in many countries’ financial sectors), and to help member countries limit regulatory arbitrage, a potential precursor for a future crisis. The Fund has also developed new analytical tools that benefit from the expanded set of financial data.

#### (iv) Use of Fund resources

##### *Data deficiencies can affect program design and monitoring . . .*

40. Data quality and availability are also extremely relevant for IMF lending. Staff must be able to count on information adequate to allow the design of a program fit for the intended purpose. This has usually been the case, but in some instances, staff has indicated that policy programs would have been formulated differently if more accurate information had been available (Reichmann and Monasterski, 2016) (Box 4). From 2000 through March 2015, there were 62 cases of misreporting vis-à-vis data in the context of Fund-supported programs,<sup>39</sup> up sharply from the nine cases in the previous 15-year period from 1985 to 2000. Occasionally, inaccurate or incomplete information about a member country’s observance of a program performance criterion may give rise to a “noncomplying purchase” and the issuance of a misreporting notification to the Executive Board.

41. Even when data allow for adequate diagnosis and formulation of policies, the specific design of performance criteria is influenced by considerations of data accuracy, availability, and timeliness. Trade-offs are unavoidable among these factors, and the resulting criteria will seldom be totally homogeneous across time

<sup>37</sup>These datasets are part of the Data Gaps Initiative.

<sup>38</sup>The improvement in the collection of FSIs is especially noteworthy, with 101 countries currently providing at least the core indicators as of mid-2015, compared with 57 in 2007. Nonetheless, FSIs notably suffer from lack of comparability across countries, as they are based on very heterogeneous definitions of capital, nonperforming loans, etc.

<sup>39</sup>Of these 62 cases of provision of incorrect data, 11 were considered “de minimis,” 38 received waivers, and only 13 required corrective actions, usually involving early repurchase or repayment.